A Weiszfeld-like algorithm for a Weber location problem constrained to a closed and convex set[☆]

Germán A. Torres[a,*]

[a] *Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, CIEM (CONICET), Medina Allende s/n, Ciudad Universitaria (5000) Córdoba, Argentina*

Abstract

The Weber problem consists of finding a point in $\mathbb{R}^n$ that minimizes the weighted sum of distances from m points in $\mathbb{R}^n$ that are not collinear. An application that motivated this problem is the optimal location of facilities in the 2-dimensional case. A classical method to solve the Weber problem, proposed by Weiszfeld in 1937, is based on a fixed point iteration.

In this work a Weber problem constrained to a closed and convex set is considered. A Weiszfeld-like algorithm, well defined even when an iterate is a vertex, is presented. The iteration function Q that defines the proposed algorithm, is based mainly on an orthogonal projection over the feasible set, combined with the iteration function of a modified Weiszfeld algorithm presented by Vardi and Zhang in 2001.

It can be seen that the proposed algorithm generates a sequence of feasible iterates that have descent properties. Under certain hypotheses, the limit of this sequence satisfies the KKT optimality conditions, is a fixed point of the iteration function that defines the algorithm, and is the solution of the constrained minimization problem. Numerical experiments confirmed the theoretical results.

Keywords: location, Weber problem, Weiszfeld algorithm, fixed point iteration
2000 MSC: 90B85, 90C25, 90C30

1. Introduction

Let $a^1, \ldots, a^m$ be m distinct points in the space $\mathbb{R}^n$, called vertices, and positive numbers $w_1, \ldots, w_m$, called weights. The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \sum_{j=1}^{m} w_j \left\| x - a^j \right\|, \tag{1}$$

is called the Weber function, where $\|\cdot\|$ denotes the Euclidean norm. It is well-known that this function is not differentiable at the vertices, and strictly convex if the vertices are not collinear (we will assume this hypothesis from now on).

The Weber problem (also known as the Fermat-Weber problem) is to find a point in $\mathbb{R}^n$ that minimizes the weighted sum of Euclidean distances from the m given points, that is, we have to find the solution of the following unconstrained optimization problem:

$$\begin{array}{rl} \underset{x}{\operatorname{argmin}} & f(x) \\ \text{subject to} & x \in \mathbb{R}^n. \end{array} \tag{2}$$

[☆] Research partially supported by the CONICET, the SECYT-UNC and the ANPCyT.

[*] Corresponding author
Email address: torres@famaf.unc.edu.ar (Germán A. Torres)

This problem has a unique solution x^u in $\mathbb{R}^n$.

The problem was also stated as a pure mathematical problem by Fermat [44, 27], Cavalieri [37], Steiner [14], Fasbender [20] and many others. Several solutions, based on geometrical arguments, were proposed by Torricelli and Simpson. In [30] historical details and geometric aspects were presented by Kupitz and Martini. In [41] Weber formulated the problem (2) from an economical point of view. The vertices represent customers or demands, the solution to the problem denotes the location of a new facility, and the weights are costs associated with the interactions between the new facility and the customers.

Among several schemes to solve the Weber location problem (see [12, 19, 28, 34]), one of the most popular methods was presented by Weiszfeld in [42, 43]. The Weiszfeld algorithm is an iterative method based on the first-order necessary conditions for a stationary point of the objective function.

If we define $T_0 : \mathbb{R}^n \to \mathbb{R}^n$ by:

$$T_0(x) = \begin{cases} \dfrac{\displaystyle\sum_{j=1}^{m} \frac{w_j a^j}{\|x - a^j\|}}{\displaystyle\sum_{j=1}^{m} \frac{w_j}{\|x - a^j\|}}, & \text{if } x \neq a^1, \dots, a^m, \\ \\ a^k, & \text{if } x = a^k,\ k = 1, \dots, m, \end{cases} \tag{3}$$

the Weiszfeld algorithm is:

$$x^{(l)} = T_0\left(x^{(l-1)}\right), \quad l \in \mathbb{N}, \tag{4}$$

where $x^{(0)} \in \mathbb{R}^n$ is a starting point.

The Weiszfeld algorithm (4), despite of its simplicity, has a serious problem if some $x^{(l)}$ lands accidentally in a vertex a^k, because the algorithm gets stuck at a^k, even when a^k is not the solution of (2). Many authors studied the set of initial points for which the sequence generated by the Weiszfeld algorithm yields in a vertex (see [29, 11, 6, 9, 7, 3]). Vardi and Zhang [40] derived a simple but nontrivial modification of the Weiszfeld algorithm in which they solved the problem of landing in a vertex.

Generalizations and new techniques for the Fermat-Weber location problem have been developed in recent years. In [18] Eckhardt applied the Weiszfeld algorithm to generalized Weber problems in Banach spaces. An exact algorithm for a Weber problem with attraction and repulsion was presented by Chen et al. in [13]. Kaplan and Yang [24] proved a duality theorem which includes as special cases a great variety of choices of norms in the terms of the Fermat-Weber sum. In [10] Carrizosa et al. studied the so called Regional Weber Problem, which allows the demand not to be concentrated onto a finite set of points, but follows an arbitrary probability measure. In [17] Drezner and Wesolowsky studied the case where different l_p norms are used for each demand point. In [23] the so called Complementary Problem (the Weber problem with one negative weight) was studied by Jalal and Krarup, and geometrical solutions were given. In [15] Drezner presented a Weiszfeld-like iterative procedure and convergence is proved if appropriate conditions hold.

In some practical problems it is necessary to consider barriers (forbidden regions). Barriers were first introduced to location modeling by Katz and Cooper [25]. There exist several heuristic and iterative algorithms for single-facility location problems for distance computations in the presence of barriers (see [2, 8, 5, 4]). In [35] Pfeiffer and Klamroth presented a unified formulation for problems with barriers and network location problems. A complete reference to barriers in location problems can be found in [26]. Barriers can be applied to model real life problems where regions like lakes and mountains are forbidden.

On the other hand, there are location problems whose solution needs to lie within a closed set. For example, see [39] for a discussion of the case when the solution is constrained to be within a

maximum distance of each demand point. Drezner and Wesolowsky [16] studied the problem of locating an obnoxious facility with rectangular distances (l_1 norm), where the facility must lie within some prespecified region (linear constraints). A primal-dual algorithm to deal with the constrained Fermat-Weber problem using mixed norms was developed in [33] by Idrissi et al.. In [21] Hansen et al. presented an algorithm for solving the Weber problem when the set of feasible locations is the union of a finite number of convex polygons. In [36] Pilotta and Torres considered a Weber location problem with box constraints.

Constrained Weber problems arise when we require that the solution is in an area (feasible region) determined by, for example, environmental and/or political reasons. It could be the case for a facility producing dangerous materials that must be installed in a restricted (constrained) area. Another example could be the location of a plant in an industrial zone or of a hospital in a non-polluted area.

In this paper a constrained location problem is considered. An algorithm is proposed to solve the following problem:

$$\begin{array}{rl} \underset{x}{\operatorname{argmin}} & f(x) \\ \text{subject to} & x \in \Omega, \end{array} \tag{5}$$

where Ω is a closed and convex set, generalizing the problem formulated in [36]. Problem (5) could be seen as a nonlinear programming problem and solved by standard solvers, but they may fail since the Weber function is not differentiable at the vertices.

It can be proved that problem (5) has a unique solution x^*, since the function f is strictly convex and Ω is a closed and convex set. On the other hand, it is well-known that the convex hull of the given vertices $a^1, \dots, a^m$ contains the solution x^u of the unconstrained Weber problem (see for instance [29, pp. 100]). If Ω contains the convex hull, both solutions x^* and x^u agree. In other cases, the solution x^* is not necessarily a projection of x^u over Ω (see [36]). The algorithm is based basically on a slight variation of an orthogonal projection of the Weiszfeld algorithm presented in [40], that is well defined even when an iterate coincides with a vertex. Properties of the sequence generated by the proposed algorithm related with the minimization problem 5 will be proved in the following sections.

The paper is structured as follows: Section 2 describes the results in [40] in which a modified Weiszfeld algorithm is presented and some notation is introduced. In Section 3 the proposed algorithm is defined. Section 4 is dedicated to definitions and technical lemmas. In Section 5 the main results about convergence to optimality are presented. Numerical experiments are considered in Section 6. Finally, conclusions are given in Section 7.

Some words about notation. As it was mentioned, we will call x^u the solution of problem (2) and x^* the solution of problem (5). The symbols $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ will refer to the standard Euclidean norm and standard inner product in $\mathbb{R}^n$, respectively. For a function $f : \mathbb{R} \to \mathbb{R}$ we will denote by $f'(a-)$ the left-hand side derivative at a, and by $f'(a+)$ the right-hand side derivative at a.

2. The modified Weiszfeld algorithm

This section reviews the main results presented in [40] in which the authors generalize the Weiszfeld algorithm for the case that an iterate lands on a vertex. From now on, this algorithm will be referred to as the modified Weiszfeld algorithm.

In order to make notation easier, we define the function $A : \mathbb{R}^n \to \mathbb{R}$ by:

$$A(x) = \begin{cases} \displaystyle\sum_{j=1}^{m} \frac{w_j}{2\,\|x - a^j\|}, & \text{if } x \neq a^1, \dots, a^m, \\[2ex] \displaystyle\sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{2\,\|a^k - a^j\|}, & \text{if } x = a^k,\ k = 1, \dots, m. \end{cases} \tag{6}$$

Notice that $A(x) > 0$ for all $x \in \mathbb{R}^n$. In [40, pp. 563], the number $A(a^k)$ was called A_k.

A generalization for the iteration function T_0, defined in (3), is given by $\widetilde{T} : \mathbb{R}^n \to \mathbb{R}^n$ defined as follows:

$$\widetilde{T}(x) = \begin{cases} \dfrac{\displaystyle\sum_{j=1}^{m} \frac{w_j a^j}{\|x - a^j\|}}{2A(x)}, & \text{if } x \neq a^1, \ldots, a^m, \\[2em] \dfrac{\displaystyle\sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j a^j}{\|a^k - a^j\|}}{2A\left(a^k\right)}, & \text{if } x = a^k,\ k = 1, \ldots, m. \end{cases} \tag{7}$$

Notice that $\widetilde{T}$ coincides with T_o in $\mathbb{R}^n - \{a^1, \ldots, a^m\}$.

Let $\widetilde{R} : \mathbb{R}^n \to \mathbb{R}^n$ and $r : \mathbb{R}^n \to \mathbb{R}$ be:

$$\widetilde{R}(x) = \begin{cases} \displaystyle\sum_{j=1}^{m} \frac{w_j\left(a^j - x\right)}{\|x - a^j\|}, & \text{if } x \neq a^1, \ldots, a^m, \\[2em] \displaystyle\sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j\left(a^j - a^k\right)}{\|a^k - a^j\|}, & \text{if } x = a^k,\ k = 1, \ldots, m, \end{cases} \tag{8}$$

$$r(x) = \|\widetilde{R}(x)\|, \qquad \forall\, x \in \mathbb{R}^n.$$

The function $\widetilde{R}$ generalizes the negative gradient of the Weber function since, for all $x \neq a^1, \ldots, a^m$,

$$\nabla f(x) = -\widetilde{R}(x). \tag{9}$$

The following lemma is very easy to prove (see [40, equation (14)]), and it relates the functionals $\widetilde{T}$ and $\widetilde{R}$.

Lemma 1. *For all $x \in \mathbb{R}^n$ we have $\widetilde{R}(x) = 2A(x)\left[\widetilde{T}(x) - x\right]$.*

If we define $\gamma : \mathbb{R}^n \to \mathbb{R}$ by:

$$\gamma(x) = \begin{cases} 0, & \text{if } x \neq a^1, \ldots, a^m, \\ 0, & \text{if } x = a^k \text{ and } r\left(a^k\right) = 0 \text{ for some } k = 1, \ldots, m, \\ w_k / r\left(a^k\right), & \text{if } x = a^k \text{ and } r\left(a^k\right) \neq 0 \text{ for some } k = 1, \ldots, m, \end{cases}$$

we can see that $\gamma(x) \geq 0$ for all $x \in \mathbb{R}^n$.

The modified Weiszfeld algorithm presented in [40] is defined by:

$$x^{(l)} = T\left(x^{(l-1)}\right), \quad l \in \mathbb{N},$$

where $x^{(0)} \in \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by:

$$T(x) = (1 - \beta(x))\, \widetilde{T}(x) + \beta(x) x, \tag{10}$$

where $\beta : \mathbb{R}^n \to \mathbb{R}$ is defined by $\beta(x) = \min\{1, \gamma(x)\}$.

Remark 2. *(a) If $x \neq a^1, \ldots, a^m$, then $\beta(x) = 0$ because $\gamma(x) = 0$. So, we can deduce that $T(x) = \widetilde{T}(x)$. Notice that this fact implies that the functional T is continuous in $\mathbb{R}^n - \{a^1, \ldots, a^m\}$.*

(b) It can be seen that if $a^k \neq x^u$, then $0 < \beta(a^k) < 1$ (see [40, pp. 563]).

(c) From equation (10) we obtain that $T(x) - x = (1-\beta(x))\left(\widetilde{T}(x) - x\right)$ for $x \in \mathbb{R}^n$.

The main result in [40, pp. 562] is:

Theorem 3. *The following propositions are equivalent:*

(a) $x = x^u$.

(b) $T(x) = x$.

(c) $r(x) \leq \eta(x)$.

where

$$\eta(x) = \begin{cases} 0, & \text{if } x \neq a^1, \ldots, a^m, \\ w_k, & \text{if } x = a^k,\ k = 1, \ldots, m. \end{cases}$$

3. The proposed algorithm

This section is dedicated to describe the proposed algorithm, introducing some definitions and remarks.

First of all, we can notice that problem (5) has a unique solution, due to the fact that f is a non-negative, strictly convex, and continuous function, $\lim_{\|x\|\to\infty} f(x) = \infty$ and Ω is closed and convex.

In order to define the proposed algorithm at the vertices, we will need to determine which points of the segment that joins a^k and $T(a^k)$ are in the feasible set Ω. If $k = 1, \ldots, m$, let the set $\mathcal{S}_k$ be defined by:

$$\mathcal{S}_k = \left\{\lambda \in [0,1] : (1-\lambda)T(a^k) + \lambda a^k \in \Omega\right\}.$$

Notice that $\mathcal{S}_k$ could be equal to the empty set in case that a^k and $T(a^k)$ do not belong to Ω. On the other hand, if $a^k \in \Omega$, then $1 \in \mathcal{S}_k$, which means that $\mathcal{S}_k \neq \emptyset$. Thus, we can define:

$$\lambda(a^k) = \inf \mathcal{S}_k, \quad a^k \in \Omega.$$

In case a vertex a^k is not in Ω, there is no need to define the number $\lambda(a^k)$.

In the following lemma, a set of basic properties of $\lambda(a^k)$ are listed:

Lemma 4. *If $k = 1, \ldots, m$ and $a^k \in \Omega$ then:*

(a) $\lambda(a^k) \in [0,1]$.

(b) If $T(a^k) \in \Omega$ then $\lambda(a^k) = 0$.

(c) If $T(a^k) \notin \Omega$ then $\lambda(a^k) \in (0,1]$.

Proof. The proof of (a) follows from the definition of $\mathcal{S}_k$. If $T(a^k) \in \Omega$, then $0 \in \mathcal{S}_k$, so $\lambda(a^k) = 0$, and this proves (b). Finally, for item (c), let us consider that $T(a^k) \notin \Omega$. Since Ω is a closed set, there is an entire ball centered at $T(a^k)$ that does not intersect Ω, which implies that there exists ϵ such that $(1-\lambda)T(a^k) + \lambda a^k \notin \Omega$ for all $\lambda \in [0, \epsilon]$. Thus, $\lambda(a^k) \in (0,1]$ and this concludes the proof. $\square$

Let us call $P_\Omega : \mathbb{R}^n \to \Omega$ the orthogonal projection over Ω. Since Ω is a nonempty, closed and convex set, the operator P_Ω is a continuous function [1, pp. 99].

We define the iteration function $Q : \Omega \to \Omega$ by:

$$Q(x) = \begin{cases} P_\Omega \circ T(x), & \text{if } x \neq a^1, \ldots, a^m, \\ \left(1 - \lambda(a^k)\right) T(a^k) + \lambda(a^k) a^k, & \text{if } x = a^k \in \Omega,\ k = 1, \ldots, m. \end{cases} \tag{11}$$

There will be no need to define Q outside Ω since the proposed algorithm generates a sequence of feasible points. The iteration function Q at $x \in \Omega$ coincides with the orthogonal projection of $T(x)$ over the feasible set when x is different from the vertices. Only when x is a vertex a^k belonging to Ω, $Q(x)$ is defined as the farthest possible feasible point of the segment that joins x with $T(x)$.

The following remark states some basic properties of the iteration function of the proposed algorithm.

Remark 5.

(a) If $a^k \in \Omega$ and $T(a^k) \in \Omega$, then $Q(a^k) = T(a^k) = P_\Omega \circ T(a^k)$.

(b) If $a^k \in \Omega$, it can be seen that:

$$\begin{aligned} Q(a^k) - a^k &= \left(1 - \lambda(a^k)\right)\left(T(a^k) - a^k\right), \\ Q(a^k) - T(a^k) &= -\lambda(a^k)\left(T(a^k) - a^k\right). \end{aligned}$$

(c) The functional Q is continuous in $\mathbb{R}^n - \{a^1, \ldots, a^m\}$.

Proof. The proofs of (a) and (b) are straightforward. For (c), since P_Ω is continuous in $\mathbb{R}^n$ (see [1, pp. 99]) and T is continuous in $\mathbb{R}^n - \{a^1, \ldots, a^m\}$ (see Remark 2), we have that Q is continuous in $\mathbb{R}^n - \{a^1, \ldots, a^m\}$. □

The proposed algorithm is described below.

Algorithm 6. *Let $\Omega \subset \mathbb{R}^n$ be a closed and convex set. Assume that $x^{(0)} \in \Omega$ is an initial approximation such that $f(x^{(0)}) \leq f(a^j)$ for all $j \in \{1, \ldots, m\}$ and $a^j \in \Omega$. Given $\varepsilon > 0$ a tolerance and $x^{(l-1)} \in \Omega$, do the following steps to compute $x^{(l)}$:*

Step 1: *Compute:*

$$x^{(l)} = Q\left(x^{(l-1)}\right). \tag{12}$$

Step 2: *Stop the execution if*

$$\left\|x^{(l)} - x^{(l-1)}\right\| < \varepsilon,$$

and declare $x^{(l)}$ as solution to the problem (5). Otherwise return to Step 1.

From the definition of Q it follows that Algorithm 6 generates a sequence of feasible iterates. Also notice that if there are vertices in the feasible set, $x^{(0)}$ can be one of them, for example, a vertex a_s such that $f(a_s) \leq f(a^j)$ for all $a^j \in \Omega$. On the other hand, if there are no vertices in the feasible set, $x^{(0)}$ can be chosen as the projection over Ω of the null vector.

4. Some definitions and technical results

The purpose of this section is to define some entities and prove technical lemmas that will be important in the proof of the main results.

First of all, we will define some useful operators for making notation easier. If $\mathcal{A} \subset \{1, \ldots, n\}$, then we define $\|\cdot\|_\mathcal{A} : \mathbb{R}^n \to \mathbb{R}$ and $\langle \cdot, \cdot \rangle_\mathcal{A} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by:

$$\|x\|_\mathcal{A} = \sqrt{\sum_{j \in \mathcal{A}} x_j^2}, \qquad \langle x, y \rangle_\mathcal{A} = \sum_{j \in \mathcal{A}} x_j y_j.$$

Notice that, when $\mathcal{A} \subsetneq \{1, \ldots, n\}$, $\|\cdot\|_\mathcal{A}$ is not necessarily a norm and $\langle \cdot, \cdot \rangle_\mathcal{A}$ is not necessarily an inner product.

According to this definition, if $\mathcal{A}$ and $\mathcal{B}$ are sets such that $\mathcal{A} \cap \mathcal{B} = \emptyset$ and $\mathcal{A} \cup \mathcal{B} = \{1, \dots, n\}$, it can be seen that:

$$\begin{aligned} \|x\|^2 &= \|x\|^2_{\mathcal{A}} + \|x\|^2_{\mathcal{B}}, && (13)\\ \langle x, y\rangle &= \langle x, y\rangle_{\mathcal{A}} + \langle x, y\rangle_{\mathcal{B}}, && (14)\\ c\langle x, y\rangle_{\mathcal{A}} &= \langle cx, y\rangle_{\mathcal{A}} = \langle x, cy\rangle_{\mathcal{A}}. && (15)\end{aligned}$$

For $x \in \Omega$, let us define the following sets of indices:

$$\begin{aligned} \mathcal{N}(x) &= \{k \in \mathbb{N} : \quad 1 \leq k \leq n, \quad (T(x))_k \neq (Q(x))_k\},\\ \mathcal{E}(x) &= \{k \in \mathbb{N} : \quad 1 \leq k \leq n, \quad (T(x))_k = (Q(x))_k\},\end{aligned}$$

Notice that for all $x \in \mathbb{R}^n$ we have that $\mathcal{N}(x) \cap \mathcal{E}(x) = \emptyset$ and $\mathcal{N}(x) \cup \mathcal{E}(x) = \{1, \dots, n\}$.

Let $\alpha : \Omega \to \mathbb{R}^n$ be the following function:

- If $x \neq a^1, \dots, a^m$:

$$\alpha(x) = \sum_{j=1}^{m} \frac{w_j}{\|x - a^j\|} \left[Q(x) - a^j\right]. \tag{16}$$

- If $x = a^k \in \Omega$ for some $k = 1, \dots, m$:

$$\alpha(x) = \sum_{\substack{j=1\\ j\neq k}}^{m} \frac{w_j}{\|a^k - a^j\|} \left[Q(a^k) - (1 - \beta(a^k))a^j - \beta(a^k)a^k\right]. \tag{17}$$

It can be seen that the function α is related to the iteration function Q of the proposed algorithm, and the iteration function T of the modified algorithm.

Lemma 7. *If $x \in \Omega$, then $\alpha(x) = 2A(x)\left[Q(x) - T(x)\right]$.*

Proof. If $x \neq a^1, \dots, a^m$, then:

$$\begin{aligned} \alpha(x) &= \sum_{j=1}^{m} \frac{w_j}{\|x - a^j\|}\left[Q(x) - a^j\right] = \sum_{j=1}^{m} \frac{w_j Q(x)}{\|x - a^j\|} - \sum_{j=1}^{m} \frac{w_j a^j}{\|x - a^j\|}\\ &= \left(\sum_{j=1}^{m} \frac{w_j}{\|x - a^j\|}\right)\left[Q(x) - \frac{\sum_{j=1}^{m} \frac{w_j a^j}{\|x-a^j\|}}{\sum_{j=1}^{m} \frac{w_j}{\|x-a^j\|}}\right] = 2A(x)\left[Q(x) - \widetilde{T}(x)\right]\\ &= 2A(x)\left[Q(x) - T(x)\right].\end{aligned}$$

where in the last equalities we have used the definition of $\widetilde{T}$ as in (7), and the fact that $\widetilde{T}(x) = T(x)$ due to Remark 2.

If $x = a^k$ for some $k = 1, \dots, m$, we follow a similar procedure than in the previous case. □

Now, we will define auxiliary functions that take into account the projection P_Ω in order to prove a descent property of f (see next sections). If $x \in \Omega$, we define:

(a) $E_x : \mathbb{R}^n \to \mathbb{R}^n$, where:

$$(E_x(y))_k = \begin{cases} (Q(x))_k, & \text{if } k \in \mathcal{N}(x),\\ y_k, & \text{if } k \in \mathcal{E}(x). \end{cases} \tag{18}$$

(b) If $\mathcal{E}(x) = \{i_1, \dots, i_r\} \neq \emptyset$ define $P_x : \mathbb{R}^n \to \mathbb{R}^r$ where:

$$(P_x(y))_k = y_{i_k}, \quad k = 1, \dots, r.$$

A useful property of E_x, that follows from the definition, is pointed out in the following remark.

Remark 8. *If $x \in \Omega$ then $E_x \circ Q(x) = Q(x)$.*

The iteration function Q inherits an important property from the orthogonal projection P_Ω.

Lemma 9. *If $x \in \Omega$ we have that $\langle Q(x) - x, Q(x) - T(x) \rangle \leq 0$.*

Proof. If $x \neq a^1, \dots, a^m$, then $Q(x) = P_\Omega \circ T(x)$. By a property of the orthogonal projection [1, pp. 93] we have that $\langle Q(x) - x, Q(x) - T(x) \rangle \leq 0$.

If $x = a^k$ for some $k = 1, \dots, m$, Remark 5 and Lemma 4 imply:

$$\left\langle Q(a^k) - a^k, Q(a^k) - T(a^k) \right\rangle = -\lambda(a^k) \left(1 - \lambda(a^k)\right) \left\| T(a^k) - a^k \right\|^2 \leq 0,$$

and this concludes the proof. □

The next technical lemma will help us to save computations in other lemmas.

Lemma 10. *If $x \in \Omega$, $\mathcal{A}$ is a subset of $\{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, then:*

$$\left\| Q(x) - a^j \right\|_{\mathcal{A}}^2 = \left\| x - a^j \right\|_{\mathcal{A}}^2 - \| Q(x) - x \|_{\mathcal{A}}^2 + 2 \left\langle Q(x) - x, Q(x) - a^j \right\rangle_{\mathcal{A}}.$$

Proof. If $x \in \Omega$, we have:

$$\begin{aligned}
\| Q(x) - a^j \|_{\mathcal{A}}^2 &= \langle Q(x) - a^j, Q(x) - a^j \rangle_{\mathcal{A}} \\
&= \langle Q(x) - x + x - a^j, Q(x) - x + x - a^j \rangle_{\mathcal{A}} \\
&= \| Q(x) - x \|_{\mathcal{A}}^2 + \| x - a^j \|_{\mathcal{A}}^2 + 2 \langle Q(x) - x, x - a^j \rangle_{\mathcal{A}} \\
&= \| Q(x) - x \|_{\mathcal{A}}^2 + \| x - a^j \|_{\mathcal{A}}^2 + 2 \langle Q(x) - x, x - Q(x) \rangle_{\mathcal{A}} \\
&+ 2 \langle Q(x) - x, Q(x) - a^j \rangle_{\mathcal{A}} \\
&= \| x - a^j \|_{\mathcal{A}}^2 - \| Q(x) - x \|_{\mathcal{A}}^2 + 2 \langle Q(x) - x, Q(x) - a^j \rangle_{\mathcal{A}}.
\end{aligned}$$

□

If $x \in \Omega$, let us define $g_x : \mathbb{R}^n \to \mathbb{R}$ by:

$$g_x(y) = \begin{cases} \displaystyle\sum_{j=1}^{m} \frac{w_j}{2\| x - a^j \|} \| E_x(y) - a^j \|^2, & \text{if } x \neq a^1, \dots, a^m, \\ \\ \displaystyle\sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{2\| a^k - a^j \|} \| y - a^j \|^2 + w_k \| y - a^k \|, & \text{if } x = a^k, \\ \\ & k = 1, \dots, m. \end{cases} \tag{19}$$

The values that g_x assumes at x and $Q(x)$ will play an important role in the proof of a property of the objective function f.

Lemma 11. *Let $x \in \Omega$ be.*

(a) If $x \neq a^1, \ldots, a^m$ then:

$$\begin{aligned} g_x(x) &= \frac{1}{2} f(x) + 2A(x) \left\langle Q(x) - x, Q(x) - T(x) \right\rangle \\ &- A(x) \left\| Q(x) - x \right\|^2_{\mathcal{N}(x)} . \end{aligned}$$

(b) If $x = a^k$ for some $k = 1, \ldots, m$, then $g_{a^k}(a^k) = \frac{1}{2} f(a^k)$.

Proof. Let us suppose that $x \neq a^1, \ldots, a^m$. By property (13) and (18), we have for $j = 1, \ldots, m$:

$$\left\| E_x(x) - a^j \right\|^2 = \left\| x - a^j \right\|^2_{\mathcal{E}(x)} + \left\| Q(x) - a^j \right\|^2_{\mathcal{N}(x)} .$$

Using Lemma 10, we can see that:

$$\begin{aligned} g_x(x) &= \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|} \Big[\left\| x - a^j \right\|^2_{\mathcal{E}(x)} + \left\| x - a^j \right\|^2_{\mathcal{N}(x)} \\ &- \left\| Q(x) - x \right\|^2_{\mathcal{N}(x)} + 2 \left\langle Q(x) - x, Q(x) - a^j \right\rangle_{\mathcal{N}(x)} \Big] . \end{aligned}$$

Due to (13), the definition of the Weber function f, the definition of A as in (6), the property (15) and the definition of α as in (16), we obtain:

$$g_x(x) = \frac{1}{2} f(x) - A(x) \|Q(x) - x\|^2_{\mathcal{N}(x)} + \langle Q(x) - x, \alpha(x) \rangle_{\mathcal{N}(x)} .$$

By Lemma 7, the fact that $(Q(x))_i = (T(x))_i$ for all $i \in \mathcal{E}(x)$ and (14), we get:

$$\begin{aligned} g_x(x) &= \frac{1}{2} f(x) - A(x) \|Q(x) - x\|^2_{\mathcal{N}(x)} \\ &+ 2A(x) \langle Q(x) - x, Q(x) - T(x) \rangle , \end{aligned}$$

which concludes the proof of (a).

Now, let us assume that $x = a^k$ for some $k = 1, \ldots, m$. Then:

$$g_{a^k}(a^k) = \sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{2\|a^k - a^j\|} \|a^k - a^j\|^2 = \frac{1}{2} \sum_{\substack{j=1 \\ j \neq k}}^{m} w_j \|a^k - a^j\| = \frac{1}{2} f(a^k).$$

This concludes the proof of (b). □

The number $g_x(Q(x))$ can be computed in the next lemma.

Lemma 12. *Let $x \in \Omega$ be.*

(a) If $x \neq a^1, \ldots, a^m$ then:

$$\begin{aligned} g_x(Q(x)) &= \frac{1}{2} f(x) + 2A(x) \left\langle Q(x) - x, Q(x) - T(x) \right\rangle \\ &- A(x) \left\| Q(x) - x \right\|^2 . \end{aligned}$$

(b) If $x = a^k$ for some $k = 1, \ldots, m$, then

$$\begin{aligned} g_{a^k}(Q(a^k)) &= \frac{1}{2} f(a^k) - A(a^k) \left\| Q(a^k) - a^k \right\|^2 \\ &+ 2A(a^k) \left\langle Q(a^k) - a^k, Q(a^k) - T(a^k) \right\rangle \\ &- 2\beta(a^k) A(a^k) \left\langle Q(a^k) - a^k, \widetilde{T}(a^k) - a^k \right\rangle \\ &+ w_k \left\| Q(a^k) - a^k \right\| . \end{aligned}$$

Proof. First, let us consider $x \neq a^1, \ldots, a^m$. Due to Remark 8 we have:

$$g_x(Q(x)) = \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|} \|Q(x) - a^j\|^2.$$

By Lemma 10 we obtain:

$$\begin{aligned} g_x(Q(x)) &= \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|} \left[\|x - a^j\|^2 - \|Q(x) - x\|^2 \right. \\ &+ \left. 2\langle Q(x) - x, Q(x) - a^j \rangle \right]. \end{aligned}$$

Due to the definition of the Weber function f, the definition of A as in (6) and the definition of α as in (16), we deduce that:

$$g_x(Q(x)) = \frac{1}{2} f(x) - A(x)\|Q(x) - x\|^2 + \langle Q(x) - x, \alpha(x) \rangle.$$

By Lemma 7 we get:

$$g_x(Q(x)) = \frac{1}{2} f(x) - A(x)\|Q(x) - x\|^2 + 2A(x)\langle Q(x) - x, Q(x) - T(x) \rangle,$$

concluding the proof of (a).

Now, consider $x = a^k$ for some $k = 1, \ldots, m$. Due to (19) we have:

$$g_{a^k}(Q(a^k)) = \sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{2\|a^k - a^j\|} \|Q(a^k) - a^j\|^2 + w_k \|Q(a^k) - a^k\|.$$

By Lemma 10, the definition of the Weber function f and the definition of A as in (6) we obtain:

$$\begin{aligned} g_{a^k}\left(Q(a^k)\right) &= \frac{1}{2} f(a^k) - A(a^k)\|Q(a^k) - a^k\|^2 \\ &+ \left\langle Q(a^k) - a^k, \sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{\|a^k - a^j\|} \left[Q(a^k) - a^j \right] \right\rangle \\ &+ w_k \|Q(a^k) - a^k\|. \end{aligned}$$

Manipulating algebraically,

$$Q(a^k) - a^j = Q(a^k) - (1 - \beta(a^k))a^j - \beta(a^k)a^k + \beta(a^k)(a^k - a^j).$$

Due to the definition of α (see (17)) and the definition of $\widetilde{R}$ (see (8)) we get:

$$\begin{aligned} g_{a^k}\left(Q(a^k)\right) &= \frac{1}{2} f(a^k) - A(a^k)\left\|Q(a^k) - a^k\right\|^2 \\ &+ \left\langle Q(a^k) - a^k, \alpha(a^k) \right\rangle - \beta(a^k)\left\langle Q(a^k) - a^k, \widetilde{R}(a^k) \right\rangle \\ &+ w_k \|Q(a^k) - a^k\|. \end{aligned}$$

By Lemma 1 and Lemma 7 we have:

$$\begin{aligned} g_{a^k}\left(Q(a^k)\right) &= \frac{1}{2} f(a^k) - A(a^k)\left\|Q(a^k) - a^k\right\|^2 \\ &+ 2A(a^k)\left\langle Q(a^k) - a^k, Q(a^k) - T(a^k) \right\rangle \\ &- 2A(a^k)\beta(a^k)\left\langle Q(a^k) - a^k, \widetilde{T}(a^k) - a^k \right\rangle + w_k \|Q(a^k) - a^k\|. \end{aligned}$$

which concludes the proof. □

The next lemma deals with the last two terms of $g_{a^k}(Q(a^k))$.

Lemma 13. *If $a^k \in \Omega$ for some $k = 1, \dots, m$, the number*

$$z = w_k \left\|Q(a^k) - a^k\right\| - 2A(a^k)\beta(a^k) \left\langle Q(a^k) - a^k, \widetilde{T}(a^k) - a^k \right\rangle,$$

is equal to zero.

Proof. If $Q(a^k) = a^k$ the result is true. So, from now on, let us consider that $Q(a^k) \neq a^k$. First, let us check that $a^k \neq x^u$. In case that $a^k = x^u$, then $T(a^k) = a^k$ by Theorem 3. Since $a^k \in \Omega$, then $T(a^k) \in \Omega$. By Remark 5 we have that $Q(a^k) = T(a^k) = a^k$ which is a contradiction.

By Remark 2, we have that $\beta(a^k) \in (0, 1)$ (since $a^k \neq x^u$) and:

$$z = w_k \left\|Q(a^k) - a^k\right\| - \frac{2A(a^k)\beta(a^k)}{1 - \beta(a^k)} \left\langle Q(a^k) - a^k, T(a^k) - a^k \right\rangle.$$

Extracting common factors, using Remarks 2 and 5, the fact that $T(a^k) \neq a^k$ (if $T(a^k) = a^k$ then $a^k = x^u$ by Theorem 3), and the fact that $\widetilde{T}(a^k) \neq a^k$ (if $\widetilde{T}(a^k) = a^k$ then $T(a^k) = a^k$ by definition (10)) we get that:

$$\begin{aligned} z &= 2A(a^k) \left\|Q(a^k) - a^k\right\| \left\|\widetilde{T}(a^k) - a^k\right\| \left[\frac{w_k}{2A(a^k)\left\|\widetilde{T}(a^k) - a^k\right\|}\right. \\ &- \left.\beta(a^k) \left\langle \frac{(1-\lambda(a^k))(T(a^k) - a^k)}{\|(1-\lambda(a^k))(T(a^k) - a^k)\|}, \frac{T(a^k) - a^k}{\|T(a^k) - a^k\|} \right\rangle\right]. \end{aligned}$$

Simplifying and using the definition of $\beta(a^k)$ we have that:

$$z = 2A(a^k) \left\|Q(a^k) - a^k\right\| \left\|\widetilde{T}(a^k) - a^k\right\| \left[\beta(a^k) - \beta(a^k)\right] = 0,$$

which concludes the proof. $\square$

The purpose of the next two lemmas is to determine a strict inequality between the functions g_x and f at suitable points. First of all, we have to prove the following result.

Lemma 14. *Let $x \in \Omega$ be such that $x \neq Q(x)$.*

(a) If $x \neq a^1, \dots, a^m$, then $g_x(Q(x)) \leq g_x(x)$. Besides that, if $\mathcal{E}(x) \neq \emptyset$ and $P_x \circ Q(x) \neq P_x(x)$, then $g_x(Q(x)) < g_x(x)$.

(b) If $x = a^k$ for some $k = 1, \dots, m$, then $g_{a^k}(Q(a^k)) < g_{a^k}(a^k)$.

Proof. If $x \neq a^1, \dots, a^m$, then $g_x(Q(x)) - g_x(x) = -A(x)\|Q(x) - x\|^2_{\mathcal{E}(x)} \leq 0$, by Lemma 11 and Lemma 12. Besides that, if $\mathcal{E}(x) \neq \emptyset$ and $P_x \circ Q(x) \neq P_x(x)$ we deduce that $\|Q(x) - x\|_{\mathcal{E}(x)} \neq 0$. Thus, $g_x(Q(x)) < g_x(x)$.

If $x = a^k$ for some $k = 1, \dots, m$, by Lemmas 11, 12 and 13 we have:

$$\begin{aligned} g_{a^k}(Q(a^k)) - g_{a^k}(a^k) &= -A(a^k) \left\|Q(a^k) - a^k\right\|^2 \\ &+ 2A(a^k) \left\langle Q(a^k) - a^k, Q(a^k) - T(a^k) \right\rangle. \end{aligned}$$

Due to Lemma 9 and the fact that $A > 0$ we obtain:

$$g_{a^k}(Q(a^k)) - g_{a^k}(a^k) \leq -A(a^k)\|Q(a^k) - a^k\|^2 < 0,$$

and the proof is finished. $\square$

Lemma 15. *Let $x \in \Omega$ be such that $x \neq Q(x)$. Then $g_x(Q(x)) < \frac{1}{2}f(x)$.*

Proof. Let us consider the case when $x \neq a^1, \ldots, a^m$. By Lemmas 9, 11 and 14 we have that:

$$\begin{aligned} g_x(Q(x)) &\leq g_x(x) = \frac{1}{2}f(x) + 2A(x)\langle Q(x) - x, Q(x) - T(x)\rangle \\ &- A(x)\|Q(x) - x\|^2_{\mathcal{N}(x)} \\ &\leq \frac{1}{2}f(x) - A(x)\|Q(x) - x\|^2_{\mathcal{N}(x)}. \end{aligned}$$

If $\mathcal{E}(x) = \emptyset$, then $\|\cdot\|_{\mathcal{N}(x)} = \|\cdot\|$. Therefore:

$$g_x(Q(x)) \leq \frac{1}{2}f(x) - A(x)\|Q(x) - x\|^2 < \frac{1}{2}f(x).$$

If $\mathcal{E}(x) \neq \emptyset$ and $P_x \circ Q(x) = P_x(x)$, then there exists an index $i \in \mathcal{N}(x)$ such that $x_i \neq (Q(x))_i$ since $x \neq Q(x)$. Thus, $\|Q(x) - x\|_{\mathcal{N}(x)} \neq 0$, which implies:

$$g_x(Q(x)) \leq \frac{1}{2}f(x) - A(x)\|Q(x) - x\|^2_{\mathcal{N}(x)} < \frac{1}{2}f(x).$$

If $\mathcal{E}(x) \neq \emptyset$ and $P_x \circ Q(x) \neq P_x(x)$, due to Lemmas 9, 11 and 14, we have that:

$$g_x(Q(x)) < g_x(x) \leq \frac{1}{2}f(x) - A(x)\|Q(x) - x\|^2_{\mathcal{N}(x)} \leq \frac{1}{2}f(x).$$

Now, when $x = a^k$ for some $k = 1, \ldots, m$, $g_{a^k}(Q(a^k)) < g_{a^k}(a^k) = \frac{1}{2}f(a^k)$ due to Lemma 11 and Lemma 14. □

The next lemma states an equality that relates the Weber function and g_x at appropriate points when $x \neq a^1, \ldots, a^m$. Besides that, this result will be crucial in the next section.

Lemma 16. *Let $x \neq a^1, \ldots, a^m$ be such that $x \in \Omega$ and $x \neq Q(x)$. Then:*

$$g_x \circ Q(x) = \frac{1}{2}f(x) + (f(Q(x)) - f(x)) + \delta, \quad \delta \geq 0.$$

Proof. Due to the definition of g_x as in (19) and Remark 8 we get that:

$$g_x \circ Q(x) = \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|}\|Q(x) - a^j\|^2.$$

Adding and subtracting $\|x - a^j\|$ we have:

$$\begin{aligned} g_x \circ Q(x) &= \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|}\left[\|x - a^j\| + \left(\|Q(x) - a^j\| - \|x - a^j\|\right)\right]^2 \\ &= \frac{1}{2}\sum_{j=1}^{m} w_j\|x - a^j\| + \sum_{j=1}^{m} w_j\left(\|Q(x) - a^j\| - \|x - a^j\|\right) \\ &+ \sum_{j=1}^{m} \frac{w_j}{2\|x - a^j\|}\left(\|Q(x) - a^j\| - \|x - a^j\|\right)^2. \end{aligned}$$

Notice that the first term of the last equality is the Weber function (divided by two), and the last term is a non-negative number, so we will define it as δ. So, using the definition of the Weber function in the middle term we obtain:

$$g_x \circ Q(x) = \frac{1}{2}f(x) + (f(Q(x)) - f(x)) + \delta.$$

□

5. Convergence to optimality results

This section states the main results about convergence of the sequence $\left\{x^{(l)}\right\}$ generated by Algorithm 6. The next theorem establishes that if a point $x \in \Omega$ is not a fixed point of the iteration function, then the function f strictly decreases at the next iterate.

Theorem 17. *Let $x \in \Omega$ be such that $x \neq Q(x)$. Then $f(Q(x)) < f(x)$.*

Proof. Let us consider that $x \neq a^1, \ldots, a^m$. By Lemma 15, we have that:

$$g_x \circ Q(x) < \frac{1}{2} f(x).$$

By Lemma 16 we get that:

$$\frac{1}{2} f(x) + f(Q(x)) - f(x) + \delta < \frac{1}{2} f(x).$$

Simplifying the last expression we obtain:

$$f(Q(x)) - f(x) + \delta < 0.$$

Finally,

$$f(Q(x)) - f(x) \leq f(Q(x)) - f(x) + \delta < 0.$$

Therefore, $f(Q(x)) < f(x)$.

Now, consider that $x = a^k$ for some $k = 1, \ldots, m$. Following a reasoning similar than in [40, pp. 564], using Lemma 14 we have that:

$$g_{a^k} \circ Q(a^k) - g_{a^k}(a^k) < 0.$$

By definition of g_{a^k} we know that:

$$\begin{aligned} g_{a^k} \circ Q(a^k) - g_{a^k}(a^k) &= w_k \left\|Q(a^k) - a^k\right\| \\ &+ \sum_{\substack{j=1 \\ j \neq k}}^{m} \frac{w_j}{2\left\|a^k - a^j\right\|} \left(\left\|Q(a^k) - a^j\right\|^2 - \left\|a^k - a^j\right\|^2 \right). \end{aligned}$$

Using the fact that $\left(a^2 - b^2\right)/(2b) \geq a - b$ for $a = \left\|Q(a^k) - a^j\right\|^2 \geq 0$ and $b = \left\|a^k - a^j\right\|^2 > 0$ we obtain that:

$$\begin{aligned} g_{a^k} \circ Q(a^k) - g_{a^k}(a^k) &\geq w_k \left\|Q(a^k) - a^k\right\| \\ &- \sum_{\substack{j=1 \\ j \neq k}}^{m} w_j \left\|a^k - a^j\right\| + \sum_{\substack{j=1 \\ j \neq k}}^{m} w_j \left\|Q(a^k) - a^j\right\|. \end{aligned}$$

Rearranging terms we deduce that:

$$\begin{aligned} 0 &> g_{a^k} \circ Q(a^k) - g_{a^k}(a^k) \\ &= \sum_{j=1}^{m} w_j \left\|Q(a^k) - a^j\right\| - \sum_{j=1}^{m} w_j \left\|a^k - a^j\right\| = f(Q(a^k)) - f(a^k), \end{aligned}$$

and the proof is complete. □

Corollary 18. *Let $\{x^{(l)}\}$ be the sequence generated by Algorithm 6. Then the sequence $\{f(x^{(l)})\}$ is not increasing. Even more, each time $x^{(l)} \neq Q(x^{(l)})$ the sequence strictly decreases at the next iterate.*

If the sequence $\{x^{(l)}\}$ generated by Algorithm 6 were not bounded, then we could choose a subsequence $\{y^{(l)}\}$ such that $y^{(l)} \to \infty$. But this implies that $f(y^{(l)}) \to \infty$, which is a contradiction since the sequence $\{f(x^{(l)})\}$ is not increasing.

Remark 19. *The sequence $\{x^{(l)}\}$ generated by Algorithm 6 is bounded. So, there exists a subsequence convergent to a point $x^* \in \Omega$. Hence, x^* is a feasible point.*

Due to the nondifferentiability of f at the vertices $a^1, \dots, a^m$, we can not use the KKT optimality conditions at a^k. Therefore, if a^k and z are in Ω, let us define $G^z_{a^k} : [0,1] \to \mathbb{R}$ by:

$$G^z_{a^k}(t) = f(a^k + t(z - a^k)).$$

If $a^k \in \Omega$, $z \in \Omega$, $t \in [0,1]$ and Ω convex, we have that $a^k + t(z - a^k) \in \Omega$. Notice that the right-hand side derivative $G^z_{a^k}(0+)$ (or the directional derivative of f in the direction of z) exists (see [22, pp. 33]). Besides that,

$$G^z_{a^k}{}'(0+) = w_k \left\|z - a^k\right\| - \left\langle \widetilde{R}(a^k), z - a^k \right\rangle. \tag{20}$$

The next lemma shows that if we are in a vertex a^k, the directional derivative of f at a^k in the direction of $Q(a^k)$ is a descent direction.

Lemma 20. *Let $a^k \in \Omega$ be such that $T(a^k) \notin \Omega$. Then:*

$$G^z_{a^k}{}'(0+) \geq G^{Q(a^k)}_{a^k}{}'(0+), \quad \forall\, z \in \left[a^k, T(a^k)\right], \tag{21}$$

where:

$$G^{Q(a^k)}_{a^k}{}'(0+) = -2\left[1 - \beta(a^k)\right] A(a^k) \left\|\widetilde{T}(a^k) - a^k\right\| \left\|Q(a^k) - a^k\right\|. \tag{22}$$

Proof. If $T(a^k) = a^k$ then $T(a^k) \in \Omega$, which is a contradiction. Besides that, if $\widetilde{T}(a^k) = a^k$, we would have that $T(a^k) = a^k$ because of (10), and again it would be a contradiction. So, we will consider $\widetilde{T}(a^k) \neq a^k$ and $T(a^k) \neq a^k$ for the rest of the proof. Since $T(a^k) \neq a^k$, then $\beta(a^k) \in (0,1)$ (see Remark 2 and Theorem 3).

Let us prove equation (22) first. Now, by (20) we can see that:

$$G^{Q(a^k)}_{a^k}{}'(0+) = w_k \left\|Q(a^k) - a^k\right\| - \left\langle \widetilde{R}(a^k), Q(a^k) - a^k \right\rangle.$$

Notice that if $Q(a^k) = a^k$, equation (22) holds. So, let us consider from now on that $Q(a^k) \neq a^k$. By using Lemma 1 we replace $\widetilde{R}(a^k)$ and get:

$$G^{Q(a^k)}_{a^k}{}'(0+) = w_k \left\|Q(a^k) - a^k\right\| - 2A(a^k) \left\langle \widetilde{T}(a^k) - a^k, Q(a^k) - a^k \right\rangle.$$

Extracting common factors and using the definition of β when it belongs to $(0,1)$ we obtain:

$$\begin{aligned} G^{Q(a^k)}_{a^k}{}'(0+) &= 2A(a^k) \left\|Q(a^k) - a^k\right\| \left\|\widetilde{T}(a^k) - a^k\right\| \Bigg[\beta(a^k) \\ &\quad - \left\langle \frac{\widetilde{T}(a^k) - a^k}{\left\|\widetilde{T}(a^k) - a^k\right\|}, \frac{Q(a^k) - a^k}{\left\|Q(a^k) - a^k\right\|} \right\rangle \Bigg]. \end{aligned}$$

By Remarks 2 and 5 the vectors $Q(a^k)-a^k$ and $\widetilde{T}(a^k)-a^k$ are parallel, so:

$$G_{a^k}^{Q(a^k)}{}'(0+) = 2A(a^k)\left\|Q(a^k)-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|\left[\beta(a^k)-1\right].$$

260 which is equivalent to (22).

Now, let us prove (21). If $z=a^k$ then $G_{a^k}^z(t)=f(a^k)$ for all $t\in[0,1]$, thus $G_{a^k+}^{a^k}{}'(0)=0$, and therefore the inequality (21) holds. So, let us assume that $z\neq a^k$ for the rest of the proof. Using (20) and due to Lemma 1 to replace $\widetilde{R}(a^k)$:

$$G_{a^k}^z{}'(0+) = w_k\left\|z-a^k\right\| - 2A(a^k)\left\langle \widetilde{T}(a^k)-a^k, z-a^k\right\rangle.$$

261 Extracting common factors:

$$\begin{aligned} G_{a^k}^z{}'(0+) &= 2A(a^k)\left\|z-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|\left[\frac{w_k}{2A(a^k)\left\|\widetilde{T}(a^k)-a^k\right\|}\right. \\ &\quad - \left.\left\langle \frac{\widetilde{T}(a^k)-a^k}{\left\|\widetilde{T}(a^k)-a^k\right\|}, \frac{z-a^k}{\|z-a^k\|}\right\rangle\right]. \end{aligned}$$

262 Using the expression for $\beta(a^k)\in(0,1)$ we obtain:

$$\begin{aligned} G_{a^k}^z{}'(0+) &= 2A(a^k)\left\|z-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|\left[\beta(a^k)\right. \\ &\quad - \left.\left\langle \frac{\widetilde{T}(a^k)-a^k}{\left\|\widetilde{T}(a^k)-a^k\right\|}, \frac{z-a^k}{\|z-a^k\|}\right\rangle\right]. \end{aligned}$$

If z belongs to the segment that joins a^k and $T(a^k)$ we have that $z-a^k$ and $\widetilde{T}(a^k)-a^k$ are parallel vectors, then:

$$G_{a^k}^z{}'(0+) \geq -2\left[1-\beta(a^k)\right]A(a^k)\left\|z-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|.$$

We can write $z=(1-\lambda)T(a^k)+\lambda a^k$ where $\lambda\in[0,1]$. Therefore:

$$G_{a^k}^z{}'(0+) \geq -2\left[1-\beta(a^k)\right]A(a^k)\,(1-\lambda)\left\|T(a^k)-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|.$$

for all $\lambda\in[0,1]$. The minimum value of the right-hand side of the last expression happens when $\lambda=\lambda(a^k)$, so:

$$G_{a^k}^z{}'(0+) \geq -2\left[1-\beta(a^k)\right]A(a^k)\left(1-\lambda(a^k)\right)\left\|T(a^k)-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|.$$

Using Remark 5 we conclude that:

$$G_{a^k}^z{}'(0+) \geq -2\left[1-\beta(a^k)\right]A(a^k)\left\|Q(a^k)-a^k\right\|\left\|\widetilde{T}(a^k)-a^k\right\|.$$

263 □

Now we will prove an equivalence that characterizes the solution of (5) in terms of the iteration function Q. Moreover, if x^* is a regular point that is not a vertex, then x^* is a KKT point.

From now on, let us consider that

$$\Omega = \{y \in \mathbb{R}^n : g(y) \leq 0, h(y) = 0\}, \tag{23}$$

where $g : \mathbb{R}^n \to \mathbb{R}^s$ is a convex function and $h : \mathbb{R}^n \to \mathbb{R}^p$ is an affine function.

Theorem 21. *Let Ω be defined as in (23) and $x \in \Omega$. Consider the following propositions:*

(a) x is a KKT point.

(b) x is the minimizer of the problem (5).

(c) $Q(x) = x$.

If $x \neq a^1, \dots, a^m$, g and h are continuously differentiable, and x is a regular point, then (a), (b) and (c) are equivalent.

If $x = a^k$ for some $k = 1, \dots, m$, then (b) implies (c).

Proof. Let $x \neq a^1, \dots, a^m$ be. Since f is strictly convex and Ω is convex, the KKT optimality conditions are necessary and sufficient. Therefore, it holds that (a) is equivalent to (b).

Now we will prove that (b) implies (c). Let us suppose that x is the minimizer of the problem (5). If x were not a fixed point of the iteration function Q, we would have that $x \neq Q(x)$, which means that $f(Q(x)) < f(x)$ by Theorem 17. This contradicts the hypothesis.

To demonstrate that (c) implies (a), we will assume that x is a fixed point of Q, that is, $x = Q(x)$. Since $Q(x) = P_\Omega \circ T(x)$, x is the solution of:

$$\begin{array}{rl} \underset{z}{\text{argmin}} & F(z) = \dfrac{1}{2}\|z - T(x)\|^2 \\ \text{subject to} & g(z) \leq 0, \\ & h(z) = 0. \end{array}$$

Since F and g are convex, h is affine, and x is a regular point, the KKT optimality conditions hold at x. That is, there exist multipliers $\{\mu_j\}_{j=1}^s$ and $\{\lambda_j\}_{j=1}^p$ such that (see [31, 38]):

$$\begin{array}{rcll} x - T(x) + \displaystyle\sum_{j=1}^{s} \mu_j \nabla g_j(x) + \sum_{j=1}^{p} \lambda_j \nabla h_j(x) & = & 0, & \\ \mu_j g_j(x) & = & 0, & j = 1, \dots, s, \\ \mu_j & \geq & 0, & j = 1, \dots, s, \\ g(x) & \leq & 0, & \\ h(x) & = & 0. & \end{array}$$

Multiplying these equations by $2A(x)$, using equation (9), Lemma 1 and Remark 2, we obtain:

$$\begin{array}{rcll} \nabla f(x) + \displaystyle\sum_{j=1}^{s} (2A(x)\mu_j) \nabla g_j(x) + \sum_{j=1}^{p} (2A(x)\lambda_j) \nabla h_j(x) & = & 0, & \\ (2A(x)\mu_j)\, g_j(x) & = & 0, & j = 1, \dots, s, \\ (2A(x)\mu_j) & \geq & 0, & j = 1, \dots, s, \\ g(x) & \leq & 0, & \\ h(x) & = & 0. & \end{array}$$

where $\{2A(x)\mu_j\}_{j=1}^s$ and $\{2A(x)\lambda_j\}_{j=1}^p$ are multipliers. Therefore, x is a KKT point of the problem (5) (see [31, 38]).

Now, let us suppose that $x = a^k$ for some $k = 1, \dots, m$. As before, if x is a minimizer of the problem (5), then $Q(a^k) = a^k$, otherwise $f(Q(a^k)) < f(a^k)$, which would be a contradiction. $\square$

6. Numerical experiments.

The purpose of this section is to discuss the efficiency and robustness of the proposed algorithm versus a solver for nonlinear programming problems.

A prototype code of Algorithm 6 was programmed in MATLAB (version R2011a) and executed in a PC running Linux OS, Intel(R) Core(TM) i7 CPU Q720, 1.60GHz.

We have considered a closed and convex set $\Omega \subset \mathbb{R}^2$ defined by the set $\Omega = \{y \in \mathbb{R}^n : g(y) \leq 0\}$, where g is given by:

$$g(x) = \begin{bmatrix} -4 - \frac{1}{8}x + \frac{7}{72}x^2 + \frac{1}{216}x^2(x-3) + y \\ \frac{4}{5}x + y - \frac{59}{10} \\ x - \frac{11}{2} \\ \frac{3}{2}x - y - \frac{35}{4} \\ x - y - \frac{13}{2} \\ -4 + \frac{1}{8}(x-1) + \frac{1}{16}(x-1)^2 + \frac{1}{32}(x-1)^2(x-3) - y \\ -\frac{1}{3}x - y - \frac{11}{3} \\ -\frac{2}{3}x - y - \frac{13}{3} \\ -4x + y - 19 \end{bmatrix}.$$

The feasible set is defined by linear and nonlinear constraints, as it can be seen in Figure 1.

We have built 1000 different experiments where for each one:

- The number of vertices was $m = 50$.
- The vertices were normally distributed random vectors, with mean equal to 0 and standard deviation equal to 10.
- The weights were uniformly distributed random positive numbers between 0 and 10.
- Tolerance was set to $\varepsilon = 0.00001$.

On one hand, each experiment was solved using Algorithm 6 and, on the other hand, it was considered as a nonlinear programming problem and solved using function $fmincon$ (see [32] and references therein). Since the Weber function (1) is not differentiable at the vertices, nonlinear programming solvers may fail.

Let $x_m(i)$ be the solution of (5) obtained by $fmincon$ in experiment i, and $f_m(i) = f(x_m(i))$. Analogously, let $x_p(i)$ be the solution of (5) obtained by Algorithm 6 in experiment i, and $f_p(i) = f(x_p(i))$. Figure 2 shows the difference between the arrays f_m and f_p. Both methods finished succesfully in all cases, however, Algorithm 6 found equal or better results for all experiments. For example, the difference $f_m - f_p$ was greater than 0.01 in 35 experiments (the maximum difference ocurred in experiment 506).

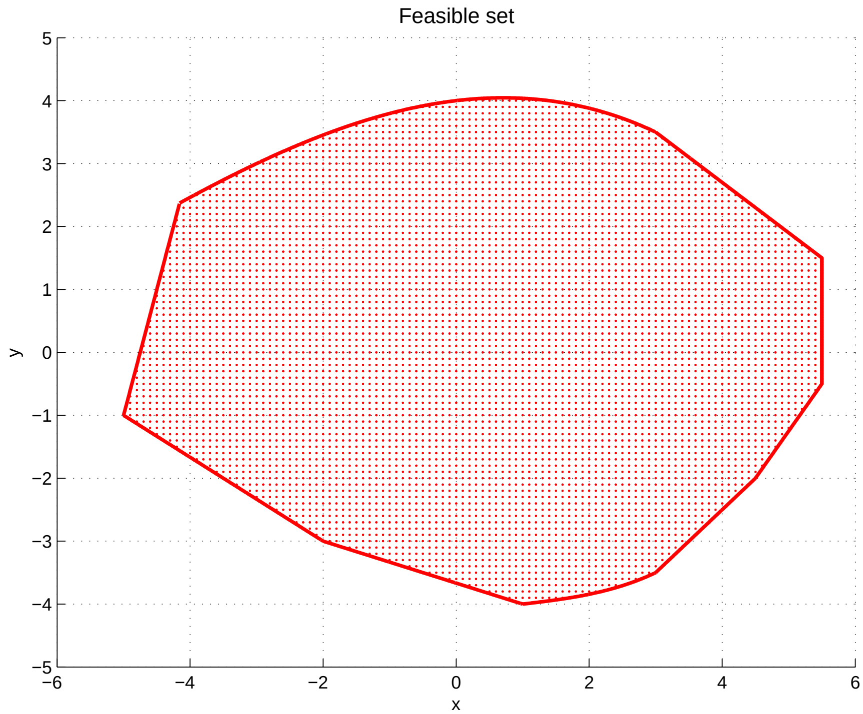

Figure 1: Feasible set Ω

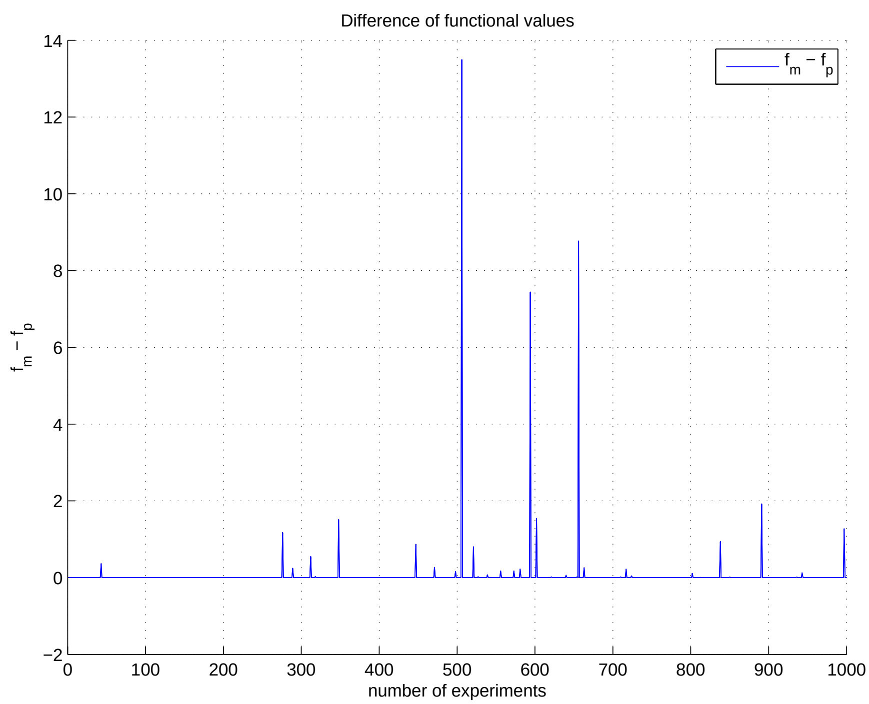

Figure 2: Difference between minimum values found by Algorithm 6 and $fmincon$.

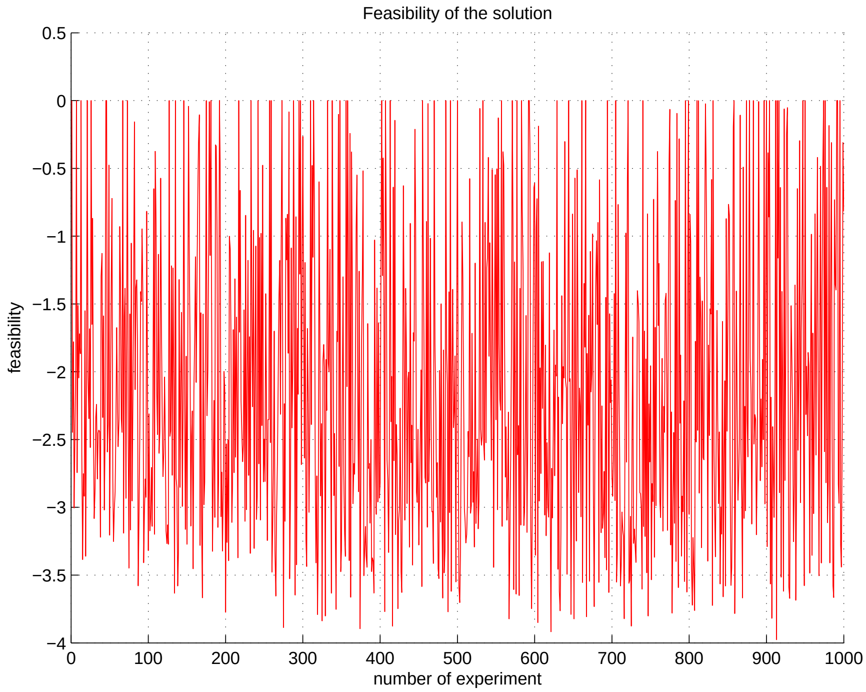

Figure 3: Feasibility of the solution $x_p(i)$ obtained by Algorithm 6.

Feasibility of the solutions $x_p(i)$ can be checked computing $\max(g(x_m(i)))$. Results can be seen in Figure 3

7. Conclusions

This paper proposes a Weiszfeld-like algorithm for solving the Weber problem constrained to a closed and convex set, and it is well defined even when an iterate is a vertex. The algorithm consists of two stages: first, iterate using the fixed point modified Weiszfeld iteration (10), and second, either project onto the set Ω when the iterate is different from the vertices, or, if the iterate is a vertex a^k, take the point belonging to the line that joins $T(a^k)$ with a^k as defined in (11).

It is proved that the constrained problem (5) has a unique solution. Besides that, the definition of the iteration function Q allows us to demonstrate that the proposed algorithm produces a sequence $\left\{x^{(l)}\right\}$ of feasible iterates. Moreover, the sequence $\left\{f\left(x^{(l)}\right)\right\}$ is not increasing, and when $x^{(l)} \neq Q\left(x^{(l)}\right)$, the sequence decreases at the next iterate. It can be seen that if a point x^* is the solution of the problem (5) then x^* is a fixed point of the iteration function Q. Even more, if x^* is different from the vertices, the fact of being x^* a fixed point of Q is equivalent to the fact that x^* satisfies the KKT optimality conditions, and equivalent to the fact that x^* is the solution of the problem (5). These properties allows us to connect the proposed algorithm with the minimization problem.

Numerical experiments showed that the proposed algorithm found equal or better solutions than a well-known standard solver, in a practical example with 1000 random choices of vertices and weights. That is due to the fact that the proposed algorithm does not use of the existence of derivatives at the vertices, because the Weber function is not differentiable at the vertices.

References

[1] N. Andréasson, A. Evgrafov, M. Patriksson, An introduction to continuous optimization, Studentlitteratur, 2007.

[2] Y.P. Aneja, M. Parlar, Technical note-Algorithms for Weber facility location in the presence of forbidden regions and/or barriers to travel, Transportation Science 28 (1994) 70–76.

[3] R.J. Cañavate Bernal, Algunas cuestiones teóricas sobre la validez del algoritmo de Weiszfeld para el problema de Weber, Rect@ Actas_13 (2005) 24.

[4] M. Bischoff, T. Fleischmann, K. Klamroth, The multi-facility location-allocation problem with polyhedral barriers, Comput. Oper. Res. 36 (2009) 1376–1392.

[5] M. Bischoff, K. Klamroth, An efficient solution method for Weber problems with barriers based on genetic algorithms, European J. Oper. Res. 177 (2007) 22–41.

[6] J. Brimberg, The Fermat-Weber location problem revisited, Mathematical Programming 71 (1995) 71–76.

[7] J. Brimberg, Further notes on convergence of the Weiszfeld algorithm, Yugosl. J. Oper. Res. 13 (2003) 199–206.

[8] S.E. Butt, T.M. Cavalier, An efficient algorithm for facility location in the presence of forbidden regions, European J. Oper. Res. 90 (1996) 56–70.

[9] L. Cánovas, R. Cañavate, A. Marín, On the convergence of the Weiszfeld algorithm, Mathematical Programming 93 (2002) 327–330.

[10] E. Carrizosa, M. Muñoz-Márquez, J. Puerto, The Weber problem with regional demand, European J. Oper. Res. 104 (1998) 358–365.

[11] R. Chandrasekaran, A. Tamir, Open questions concerning Weiszfeld's algorithm for the Fermat-Weber location problem, Mathematical Programming 44 (1989) 293–295.

[12] J.A. Chatelon, D.W. Hearn, T.J. Lowe, A subgradient algorithm for certain minimax and minisum problems, Mathematical Programming 15 (1978) 130–145.

[13] P.C. Chen, P. Hansen, B. Jaumard, H. Tuy, Weber's problem with attraction and repulsion, Journal of Regional Science 32 (1992) 467–486.

[14] R. Courant, D. Hilbert, Methoden der mathematischen Physik. I, Springer-Verlag, Berlin, 1968. Dritte Auflage, Heidelberger Taschenbucher, Band 30.

[15] Z. Drezner, On the convergence of the generalized Weiszfeld algorithm, Ann. Oper. Res. 167 (2009) 327–336.

[16] Z. Drezner, G.O. Wesolowsky, The location of an obnoxious facility with rectangular distances, Journal of Regional Science 23 (1983) 241–248.

[17] Z. Drezner, G.O. Wesolowsky, Sensitivity analysis to the value of p of the l_p distance Weber problem, Ann. Oper. Res. 111 (2002) 135–150.

[18] U. Eckhardt, Weber's problem and Weiszfeld's algorithm in general spaces, Mathematical Programming 18 (1980) 186–196.

[19] J.W. Eyster, J.A. White, W.W. Wierwille, On solving multifacility location problems using a hyperboloid approximation procedure, A I I E Transactions 5 (1973) 01–06.

[20] E. Fasbender, Über die gleichseitigen Dreiecke, welche um ein gegebenes Dreieck gelegt werden können, Journal für die reine und angewandte Mathematik 30 (1846) 230–231.

[21] P. Hansen, D. Peeters, J. Thisse, An algorithm for a constrained Weber problem, Management Science (1982) 1285–1295.

[22] J. Jahn, Introduction to the theory of nonlinear optimization, Springer Verlag, 2007.

[23] G. Jalal, J. Krarup, Geometrical solution to the Fermat problem with arbitrary weights, Ann. Oper. Res. 123 (2003) 67–104.

[24] W. Kaplan, W. Yang, Duality theorem for a generalized Fermat-Weber problem, Mathematical Programming 76 (1997) 285–297.

[25] I.N. Katz, L. Cooper, Facility location in the presence of forbidden regions. I. Formulation and the case of Euclidean distance with one forbidden circle, European J. Oper. Res. 6 (1981) 166–173.

[26] K. Klamroth, Single-facility location problems with barriers, Springer Series in Operations Research, Springer-Verlag, New York, 2002.

[27] J. Krarup, S. Vajda, On Torricelli's geometrical solution to a problem of Fermat, IMA J. Math. Appl. Bus. Indust. 8 (1997) 215–224. Duality in practice.

[28] H.W. Kuhn, On a pair of dual nonlinear programs, in: Nonlinear Programming (NATO Summer School, Menton, 1964), North-Holland, Amsterdam, 1967, pp. 37–54.

[29] H.W. Kuhn, A note on Fermat's problem, Mathematical Programming 4 (1973) 98–107.

[30] Y.S. Kupitz, H. Martini, Geometric aspects of the generalized Fermat-Torricelli problem, in: Intuitive geometry (Budapest, 1995), volume 6 of *Bolyai Soc. Math. Stud.*, János Bolyai Math. Soc., Budapest, 1997, pp. 55–127.

[31] D.G. Luenberger, Y. Ye, Linear and nonlinear programming, International Series in Operations Research & Management Science, 116, Springer, New York, third edition, 2008.

[32] MATLAB, http://www.mathworks.com/help/toolbox/optim/ug/fmincon.html, 2012.

[33] C. Michelot, O. Lefebvre, A primal-dual algorithm for the Fermat-Weber problem involving mixed gauges, Mathematical Programming 39 (1987) 319–335.

[34] M.L. Overton, A quadratically convergent method for minimizing a sum of Euclidean norms, Mathematical Programming 27 (1983) 34–63.

[35] B. Pfeiffer, K. Klamroth, A unified model for Weber problems with continuous and network distances, Comput. Oper. Res. 35 (2008) 312–326.

[36] E.A. Pilotta, G.A. Torres, A projected Weiszfeld algorithm for the box-constrained Weber location problem, Applied Mathematics and Computation 218 (2011) 2932–2943.

[37] G. Polya, Mathematik und plausibles Schließen, Band 1, Induktion Und Analogie in Der Mathematik, Birkhauser, Basel, 1962.

[38] A. Quarteroni, R. Sacco, F. Saleri, Numerical mathematics, volume 37, Springer, 2006.

[39] M.K. Schaefer, A.P.J. Hurter, An algorithm for the solution of a location problem with metric constraints, Naval Res. Logist. Quart. 21 (1974) 625–636.

[40] Y. Vardi, C.H. Zhang, A modified Weiszfeld algorithm for the Fermat-Weber location problem, Mathematical Programming 90 (2001) 559–566.

[41] A. Weber, Über den Standort der Industrien, University of Chicago Press, Chicago, 1929. Translated by C. J. Friederich: "Theory of the location of industries".

[42] E. Weiszfeld, Sur le point par lequel la somme des distances de n points donnés est minimum, Tohoku Mathematics Journal 43 (1937) 355–386.

[43] E. Weiszfeld, On the point for which the sum of the distances to n given points is minimum, Ann. Oper. Res. 167 (2009) 7–41.

[44] G.O. Wesolowsky, The Weber problem: history and perspectives, Location Sci. 1 (1993) 5–23.